

PLATE TECTONICS AND CONVECTION IN THE EARTH'S MANTLE: TOWARD A NUMERICAL SIMULATION

Numerical models of mantle convection are starting to reproduce many of the essential features of continental drift and plate tectonics. The authors show how such methods can integrate a wide variety of geophysical and geological observations.

Plate tectonics is a kinematic description of Earth that treats the outer shell of its mantle as a number of plates or rigid spherical caps that move with respect to each other (see the “Plate tectonics” sidebar). The mantle is the outer, solid 3,000-km-thick shell that overlies Earth’s fluid outer core. An enormous amount of geological and geophysical data has gone into determining the motion of the plates,¹ and within the last few years direct GPS measurements have corroborated the geological constraints on the motions of plates.

A fundamental question in geology has been, what drives the plates? This question has largely been solved—the plates are part of a system of large-scale thermal convection—and geodynamicists have moved on to more difficult questions, such as what are the details of the coupling between surface motions and deeper mantle flow? and why do we have plate tectonics as op-

posed to some other mode of tectonics or thermal convection?

Answering these more subtle questions is complicated by the fact that the primary effect of plate motion is to consume the old ocean floor and recycle it into the mantle. The primary evidence of plate history is therefore limited to the past 100 to 200 million years or so (less than 10% of the overall history of plate tectonics). We therefore rely heavily on evidence drawn from theoretical and computational models and from the continents that are not consumed wholesale by plate motions. Ideally these two sources of evidence go hand in hand to reinforce each other.

Model formulation

The equations we solve are the standard ones for thermal convection in a fluid where viscosity is high enough for inertia to be ignored:

$$\eta \nabla^2 \mathbf{u} = g \rho_0 \alpha T + \nabla p \quad (1)$$

$$\nabla \cdot \mathbf{u} = 0. \quad (2)$$

Equation 1 is the equation of motion relating the fluid velocity, \mathbf{u} , to the gravitational acceleration multiplied by the fluid density variation because of temperature, $\rho_0 \alpha T$, and the pressure gradients. The coefficient α is the coefficient of thermal expansion. The reference density,

1521-9615/00/\$10.00 © 2000 IEEE

LOUIS MORESI

Commonwealth Scientific and Industrial Research Organization

MICHAEL GURNIS

California Institute of Technology

SHIJIE ZHONG

Massachusetts Institute of Technology

Plate tectonics

Geodynamically, oceanic plates represent the top thermal boundary layer of a system of thermal convection called *mantle convection*, and this boundary layer sinks, or subducts, into the mantle at converging margins. The motion of plates—their speed and direction—is a balance between the buoyancy of cooling and thickening oceanic lithosphere and subducted slabs on the one hand and viscous flow on the other. Plate motion and the associated mantle flow appear to be a simple mode of thermal convection and have been described as a fluid-dynamical process since the seminal paper by Don Turcotte and Ron Oxburgh.¹

The process, however, is not solely fluid dynamical. Oceanic plates are mostly rigid, with little deformation within their interior, and an appreciable amount of plate motion occurs by strike-slip. Normally, in buoyancy-driven fluid flow, there would only be diverging and converging surface motions (poloidal flow), but the motions on the Earth's surface have nearly an equal amount of toroidal motion—the most significant part of the toroidal flow being strike-slip movement at plate boundaries. A substantial fraction of the total dissipation associated with plate motion could occur in the bending of the oceanic lithosphere as it subducts.² Why Earth has plate tectonics is a significant unanswered question.

Ultimately, why a planet has a particular brand of tectonics resides not only in obvious factors such as total mass and surface temperature, but in its rheology as well. Silicates are the principal building blocks of the solid Earth, which has nonlinear and temperature-, pressure-, grain-size-, and volatility-dependent rheologies. Perhaps the greatest challenge facing the development of realistic computer models of mantle convection—realistic enough to test against the rich array of geological and geophysical observations—is the incorporation of these complex rheologies.

Qualitatively, one solution to the existence of plate tectonics lies in a balance between thermal convection, the thermally activated rheology of silicates, and brittle failure of rocks at low pressures and temperatures. Stresses in the plates are largest at converging margins above downwelling mantle (subducting slabs) so that an otherwise cold and strong lithosphere fails and gives rise to weak plate margins. With this model, the lithosphere rapidly fails on geological time scales under the action of tectonic

stress. Because the rheology used in this conceptual model does not consider the material's deformational history, we refer to this as the *instantaneous rheology model*.

Clearly this is one fundamental component generating plate tectonics, but the mechanical memory of the crust and lithosphere also plays a fundamental role. An important piece of geological evidence pointing toward this alternative hypothesis is that preexisting faults and long-lived zones of preexisting weakness control the location of plate margins. The convecting system reuses old, weak structures because less energy is expended in reactivating a preexisting structure than in creating an entirely new plate margin from pristine, intact lithosphere.

An important goal of geophysics is the formulation of dynamically self-consistent, time-dependent models of mantle convection in which plate tectonics naturally arises. This is an important long-term objective. Resolving this problem will help explain why Earth has plate tectonics, while other planets, particularly Venus, do not. It will also lead to the development of dynamic models that integrate a wide variety of geological, geophysical, and geochemical observations.

Little serious debate remains concerning the dynamics controlling first-order features of instantaneous plate kinematics. The field has now moved on to equally important but unsolved problems in mantle dynamics:

- Why do we have plate tectonics, as opposed to some other tectonic mode that would remove planetary heat?
- What controls the time dependence of mantle convection?
- What is the connection between mantle dynamics and the wide range of observations against which convection models have not traditionally been compared?

The development and use of computational models are fundamental components of this endeavor.

References

1. D.L. Turcotte and E.R. Oxburgh, "Finite Amplitude Convective Cells and Continental Drift," *J. Fluid Mechanics*, Vol. 28, 1967, pp. 29–42.
2. C.P. Conrad and B.H. Hager, "Effects of Plate Bending and Fault Strength at Subduction Zones on Plate Dynamics," *J. Geophysical Research*, Vol. 104, No. B8, Aug. 1999, pp. 17551–17571.

ρ_0 , can change if the rocks undergo a phase change from one crystal form to another in response to increasing pressure and temperature within the planet. The dynamic viscosity, η , for rocks at around 1,300 degrees centigrade is enormous and similar to that of window glass at room temperature.

For simplicity's sake, these equations assume that the viscosity is constant. In practice, however, viscosity is a very strong function of temperature, pressure, composition, crystal grain size, and stress, and this cannot be ignored in computer simulations. The stress dependence is particularly important from the computational point of view

because the equations then become nonlinear. The effect is that the viscosity becomes a function of itself and we must treat the problem quite differently. The continuity equation, Equation 2, ensures conservation of mass. In this case, the constraint is stronger; it also enforces incompressibility on the flow. This equation is tightly coupled to the equation of motion, and the two must be solved as part of the same procedure.

A third equation describes the evolution through time of the temperature patterns in the fluid:

$$\frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T = Q + \kappa \nabla^2 T \quad (3)$$

Here t is time, κ is a thermal diffusivity, and Q is a heat source term, which in Earth is largely associated with the energy radioactive decay liberated. This energy equation is coupled to the equation of motion through the fact that the density of rocks changes with their temperature, and it is this density variation that drives the motion. The moving fluid carries heat with it (*advection*, which is accomplished by the $\mathbf{u} \cdot \nabla T$ term), but heat also diffuses independently of the fluid motion (the $\nabla^2 T$ term). Much of the interesting physics occurs when the advection and diffusion of heat compete in different directions and strike a dynamic balance. The energy equation contains time derivatives that the equation of motion lacks, so it is not surprising that these equations require entirely different computational methods.

Computational approach

We use a finite-element method to solve the mathematical model outlined earlier. The FEM is, in general, robust and accurate when dealing with strong variations of material properties from place to place, such as we expect to find in our situation. In part, this comes about because the equations are first integrated to an equivalent weak—or variational—form before being discretized using the mesh.

Flat Earth?

The essence of good modeling is to incorporate the essential features of the system in question while remaining as simple as possible—to make the models easier to interpret and faster to calculate. One of the traditional simplifications in mantle dynamics has been to work in a “flat-Earth” model and then to limit the computation to two dimensions (one horizontal and one vertical). This can prove very limiting when trying to draw observed behaviors of the real Earth to-

gether with the modeled internal dynamics.

Fortunately, the FEM is also well-suited to problems in which the geometry might be complicated—for example, if we need an irregular or complex mesh—this can be handled quite naturally within the standard formulation of the FEM. Using a reasonably structured mesh, the curvature of Earth can be included into models with very little additional computational overhead. Three-dimensional finite-element models remain time-consuming, and finding optimal solution algorithms requires considerable effort.

Computing the fluid velocity

In most dynamic systems, we can formulate the FEM in a time-explicit manner that is robust and simple to implement. The alternative is to use implicit methods that are more elaborate and often more temperamental but that cover much larger time increments at each step. In our case, however, the fact that inertia is negligible leaves the equation of motion independent of time. It can only be solved implicitly.

The traditional implicit approach in FEM has been to build the matrix equation and solve it directly using a method such as Crout elimination. Exploiting the fact that the finite-element matrices are quite tightly banded, we can do this relatively efficiently. Unfortunately, direct-solution methods are limited to “small” problems because the solution time scales vary rapidly with the number of unknowns (in the worst case as N^3 , where N is the number of unknowns, and even in the best case as N^2). Iterative methods can achieve much better performance than this once the problems start to become larger. For example, preconditioned conjugate-gradient methods can obtain a solution to a given accuracy in a time proportional to $N \log N$. In theory, the optimal method for our problem is multigrid, which when properly formulated can find a solution in a time proportional to N .

The multigrid method works by formulating the finite-element problem on several different scales²—usually a set of grids that are nested one within the other, sharing common nodes. The solution progresses on all of the grids at the same time, with each grid eliminating errors at a different scale. The effect is to propagate information very rapidly between different nodes in the grid that the local support of the element shape functions would otherwise prevent. In fact, a single traverse from fine to coarse grid and back can directly connect all nodes in the mesh to every other—allowing nodes that are physically

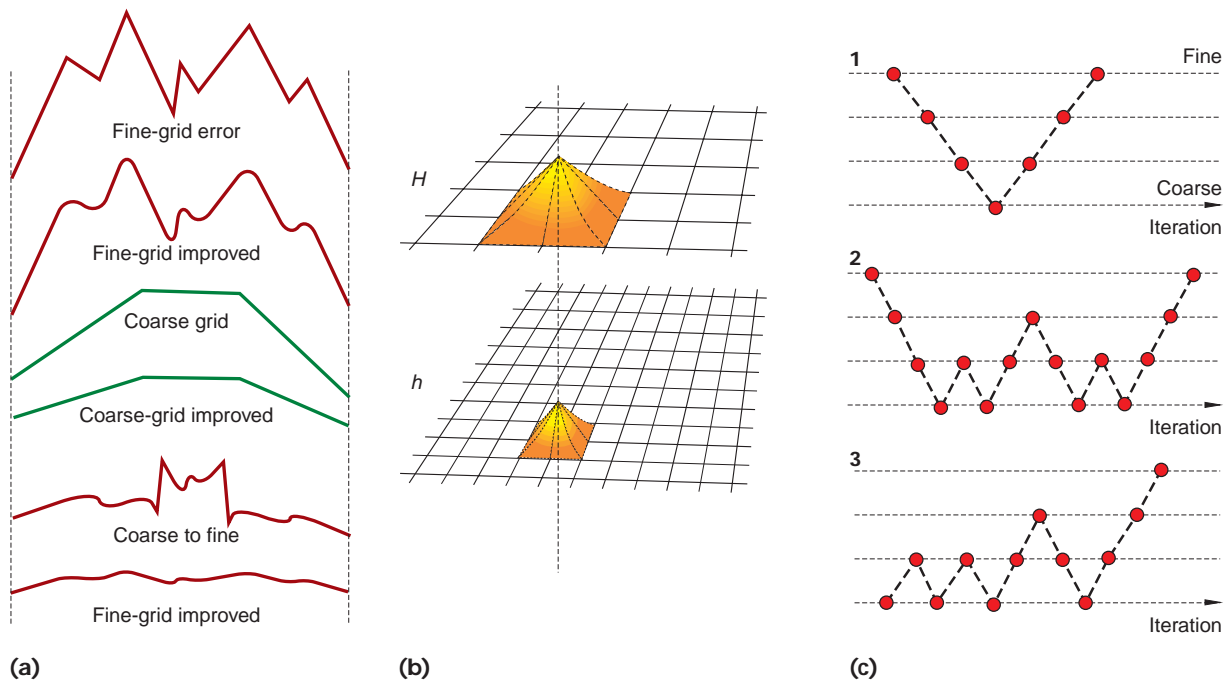


Figure 1. Multigrid solution methods: (a) Different grids can reduce errors; the fine grid reduces the short-wavelength error, whereas the coarse grid reduces the long-wavelength error. Correctly combining the two can reduce the errors at both scales in one cycle. (b) Shape functions associated with a node common to the coarse mesh— H —and the fine mesh— h —show the way in which information from different distances can be brought to a given node using different mesh scales. (c) We can devise different schemes for traversing the different meshes: (1) the V cycle, (2) the W cycle, and (3) the full multigrid method.

coupled but remote in the mesh to communicate directly during each iteration cycle.

The multigrid effect relies on using an iterative solver on each of the grid resolutions, which acts like a smoother on the residual error at the characteristic scale of that particular grid (see Figure 1a). We use Gauss-Seidel iteration because it has exactly this property. On the coarsest grid, we can use a direct solver because the number of elements is usually very small.

Incompressibility

The multigrid method leads us very quickly to a fluid velocity, but this might not, in general, be incompressible. Constraining the multigrid method to finding only incompressible solutions is a major difficulty. When using variational methods, the usual way to apply a constraint is to penalize the weak form of the equations. If there is some property we wish to make vanish (here, any compressible component of the velocity field), we multiply it by a large number and find a way to add it back into the equation we want to solve. By scaling the small constraint term, we can make it as important as the other terms.

The easiest way to see how this works for our

problem is first to write equations for a compressible flow. If the continuity equation (Equation 2) is written as

$$\nabla \cdot \mathbf{u} - p/\lambda = 0 \quad (4)$$

then in the limit that $\lambda \rightarrow \infty$, the fluid becomes incompressible. Equation 4 can be substituted into the equation of motion to give

$$\eta \nabla^2 \mathbf{u} - \lambda \nabla \nabla \cdot \mathbf{u} = \rho g \mathbf{z}. \quad (5)$$

In this way, we can make the otherwise small $\nabla \cdot \mathbf{u}$ term become dominant when λ becomes large. We have also succeeded in eliminating p as an independent variable.

This can be very effective if direct-solution methods are used, but for iterative methods it generally spells disaster. In our experience, λ of 100 or 1,000 is the maximum that can be used before the penalty terms ruin the convergence of the iteration method. This is far from the value of 10^7 to 10^9 required to suppress compressibility to the level of machine precision.

A better alternative when using an iterative solver is not to eliminate the pressure param-

ter, p , but to solve it as an unknown on the mesh. When velocity and pressure variables are combined in the same system of equations, the approach is known as a *mixed method*. Numerous strategies exist for solving the mixed formulation, but we have found only one approach that is robust enough to use with strongly varying material properties while still letting us retain the speed of the multigrid velocity solver.

This is the *Uzawa scheme*, in which an outer iteration for pressure corrects a nested set of velocity iterations that build up to an incompressible flow solution. We use a conjugate-gradient scheme for this outer iteration. The nature of the Uzawa iteration does not let us develop a reliable multigrid version of the outer iteration, but by using multigrid for the inner (velocity) iteration, the computation time still scales linearly with the number of node points in the mesh.

Nonlinear rheology

Stress-dependent viscosity makes the equation of motion nonlinear. Suddenly, it becomes necessary to introduce yet another layer of iteration to determine the viscosity at each node as a function of the changing velocity field. A triply nested iteration loop poses the danger of unmanageably long solution times. However, we have found that using multigrid concepts can again help us avoid wasting CPU time.

The viscosity is a function of the stress-field that is distributed across the mesh and is, in turn, a function of the viscosity pattern. At the start of the solution, we don't know what the ultimate viscosity field will look like. For this reason, there seems little point in using the finest grid resolution available. Instead, we operate on one of the coarser grids, using a suitably smoothed density field as the driving term. At this resolution, even the triply nested loop is not very time-consuming, and we quickly home in on the magnitude of the viscosity and stresses everywhere. We can then interpolate this information to a finer grid where we need far fewer iterations to adjust the viscosity and stresses.

This method is very close to what is known as the *full multigrid algorithm*, although the multiple nesting of iterative loops is slightly unusual.

Time and temperature

We almost always want to see the solution evolve through time in our simulations, so we generally solve this equation explicitly in time. Essentially this involves integrating Equation 3 forward through time from a given initial condition. This is generally straightforward except

for the advection term $\mathbf{u} \cdot \nabla T$.

Pure advection is a translation of the local temperature field from one part of the mesh to another: In a Lagrangian reference frame (one moving with the fluid), there is no change in temperature at any given point of reference, although the coordinate system rapidly becomes tangled. In an Eulerian reference frame (one fixed in space), the temperature at a given node point must be updated as fluid moves past it. The fluid element that is now on a particular node was, in the previous time step, at some general point in the mesh that was almost certainly not associated with a node point. The advection operation therefore requires the equivalent of an interpolation operation to determine this earlier off-node temperature. After several time steps, repeated interpolation operations start to smooth the temperature field as information is lost. The smoothing is akin to a diffusion term, so provided there is a genuine physical diffusion that operates more rapidly than the erroneous diffusion of the advection operator, it is possible to solve the problem on an Eulerian mesh.

If no diffusion occurs, or the physical diffusion coefficient is smaller than the numerical one, then we need special measures. An example might be tracking a dye (or other chemically distinct entity such as continental crust embedded in the lithosphere). We often handle this issue by dealing not with mesh-based variables but with a set of Lagrangian points that we usually refer to as tracer particles. These particles are passively moved with the flow using a Runge-Kutta integration scheme that operates over the course of a time step. An interpolation step occurs in obtaining nodal values from the particles, but the smoothing is not cumulative because this procedure simply samples—rather than disturbs—the particles.

Another problem arises from the fluid's motion relative to the grid. Although a particular discretization might be very accurate based on the temperature and velocity distribution at the beginning of a time step, the point at which the new value is computed has moved slightly with the fluid by the end of the time step. *Upwinding* is the standard cure for this problem—which is using a discretization scheme that is weighted more strongly in the upstream direction than in the downstream direction to compensate for the motion during a finite time step.

Convection models and plate tectonics

The earliest computational solutions to the

2000 Editorial Calendar



JULY/AUGUST: Humanoid Robotics

Mark Swinson, mswinson@darpa.mil, DARPA
Mobile Robots pose a unique set of challenges to artificial intelligence researchers, including issues of autonomy, uncertainty (both sensing and control), and reliability, all constrained by the discipline imposed by the physical world.

SEPTEMBER/OCTOBER: Intelligent Environments

Paddy Nixon, paddy.nixon@cs.tcd.ie, Trinity College Dublin
Smart buildings and spaces promise to give greater access and usability to both able bodied and disabled people at work or home. This issue will explore the interactions between many heterogeneous systems—between systems and networks, and between people and systems.

NOVEMBER/DECEMBER: Bringing Robots to the Classroom

Robin Murphy, murphy@csee.usf.edu, University of South Florida

How will we educate the next generation of engineers and scientists about and with robots? This issue will report educational experiences with robots and provide recommendations for further efforts.

To submit an article, visit

computer.org/intelligent/author.htm for author guidelines



GEOPHYSICS

Earth - Oceans - Atmosphere - Space - Planets
Great science at a great value

Journals

Highly cited research in print
Online for members in 2000

Books

Priced for personal libraries
New release: *Inverse Methods in Global Biogeochemical Cycles*,
Prasad Kasibhatla, et al, *Editors*

Meetings

Premier presentations in Earth systems science
2000 Spring Meeting
2000 Western Pacific Geophysics Meeting
AGU 2000 Fall Meeting

REGULAR MEMBERSHIP \$20. BELONG TO AGU TODAY.

www.agu.org

Phone: 1 800 966 2481 (toll free North America)

+1 202 462 6900 Fax: +1 202 328 0566

E-mail: orders@agu.org

Stokes convection equations made the simplifying assumption that the viscosity was constant. Despite the experimental evidence that suggests viscosity variations should dominate in the mantle, agreement with some important observations was remarkably good.

The simulations could not produce platelike motions at the surface (instead producing smoothly distributed deformation), but the average velocity, the heat flow, and the observed pattern of subsidence of the ocean floor were well matched.

More sophisticated models included the effect of temperature-dependent viscosity as a step toward more realistic simulations. In fact, Uli Christensen observed the opposite: Convection with temperature-dependent viscosity is a much worse description of the oceanic lithosphere than constant viscosity convection.³

Theoretical studies of the asymptotic limit of convection in which the viscosity variation becomes very large (comparable to values determined for mantle rocks in laboratory experiments) find that the upper surface becomes

entirely stagnant, with little or no observable motion. Vigorous convection continues underneath the stagnant layer with very little surface manifestation. This theoretical work demonstrates that the numerical simulations are producing correct results and suggests that we should look for physics beyond pure viscous flow in explaining plate motions.

The obvious association of plate boundaries with earthquake activity suggests that we'll find relevant effects in the brittle nature of the cold plates. Brittle materials have a finite strength, and if they are stressed beyond that point, they break. This is a familiar enough property of everyday materials, but rocks in the lithosphere are nonuniform, subject to great confining pressures and high temperatures, and they deform over extremely long periods. This makes it difficult to know how to apply laboratory results for rock breakage experiments to plate simulations.

An ideal, very general, rheological model for the brittle lithosphere would incorporate the effects that result from small-scale cracks, faults,

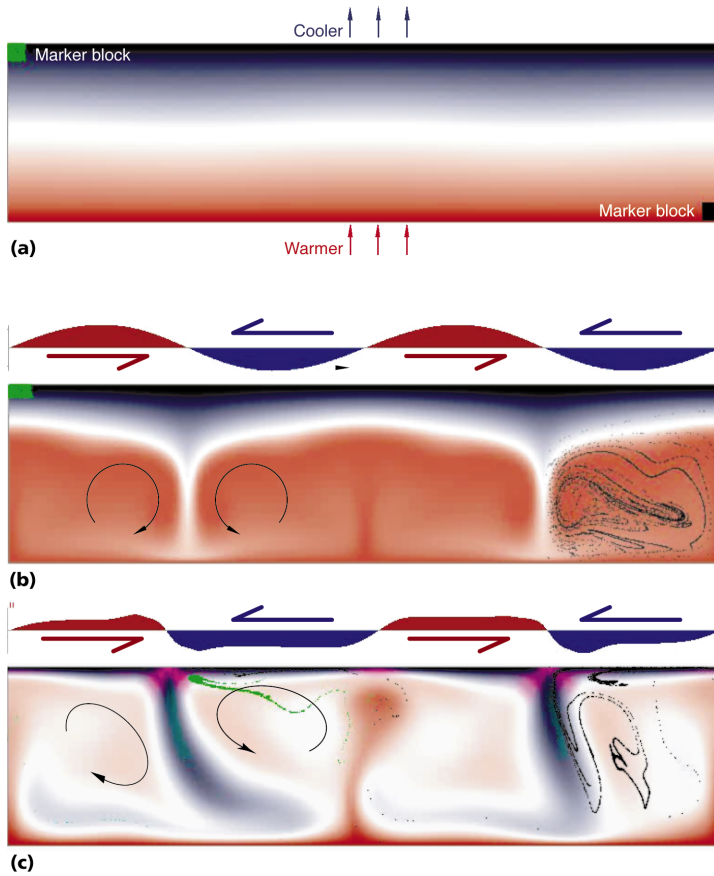


Figure 2. A layer of fluid whose viscosity decreases strongly with temperature is (a) heated from below. Initially, the temperature is set to increase uniformly from top to bottom with a very mild perturbation. Two blocks of marker particles are embedded in the fluid—green at the top, black at the bottom. (b) Stagnant lid convection develops when the yield stress is higher than typical stresses that motion in the fluid generates (the green block of particles is virtually undisturbed). Surface motion is smooth and of low magnitude—shown by the graph and the horizontal arrows. The curved arrows show the general sense of the circulation. (c) Mobile-lid convection develops when the yield stress is low enough to be overcome by convection. The surface motion is two orders of magnitude larger than in Figure 2b. The hazy purple areas show where the yield stress is exceeded. The more intense the color, the faster the material is deforming.

anisotropy, ductile shear localization caused by dynamic recrystallization, and so on. To date, most attempts to account for the brittle nature of the plates have greatly simplified the picture. Some models have imposed weak zones that represent plate boundaries; others have included sharp discontinuities that represent the plate-bounding faults; and still others have used continuum methods in which the lithosphere's yield properties are known but not the geometry of any breaks.

Of these approaches, the continuum approach is best able to demonstrate the spectrum of behaviors as convection in the mantle interacts with brittle lithospheric plates. For studying the evolution of individual plate boundaries and the fine scale of observables such as gravity and surface deformation, methods that explicitly include discontinuities work best.

The simplest possible continuum formulation includes a yield stress, τ_{yield} , expressed as a non-linear effective viscosity,

$$\eta_{eff} = \frac{\tau_{yield}}{\dot{\epsilon}} \quad (6)$$

where $\dot{\epsilon}$ is the strain rate. This formulation can be incorporated very easily into the mantle-dynamics modeling approach that we outlined because it involves making modifications only to the viscosity law. There might be some numerical difficulties, however, because the strongly non-linear rheology can lead to dramatic variations in the viscosity across relatively narrow zones.

Figure 2 compares two convecting layers with strongly temperature-dependent viscosity, one having a yield stress specified for cold material. Both simulations were started from the same slightly perturbed conductive temperature gradient, and we embedded two blocks of passive tracers in the fluids to show the deformation's evolution.

If it cannot yield, the cold boundary layer develops a stagnant lid. This is shown by the light-green block of tracers, which barely deform during the entire simulation (whereas the black block of tracers is unrecognizably distorted). The surface velocity is both small and smoothly varying. There is a slight wobble in the pattern of upwellings and downwellings as time progresses, but the basic pattern persists uninterrupted throughout.

By contrast, when the cold boundary layer can yield, no stagnant lid forms and the upper surface is highly mobile.⁴ This confines the deformation at the surface to relatively narrow zones of divergence and convergence. The convection pattern is very dynamic, and, although the same number of downwellings and upwellings persists through the entire lifetime of the simulation, the shapes and locations of upwellings and downwellings are constantly shifting.

An obvious limitation of this formulation is that it does not track the history of deformation with the fluid as it moves: Each parcel of fluid is considered pristine at the beginning of each time step. This means that the damage that occurs at

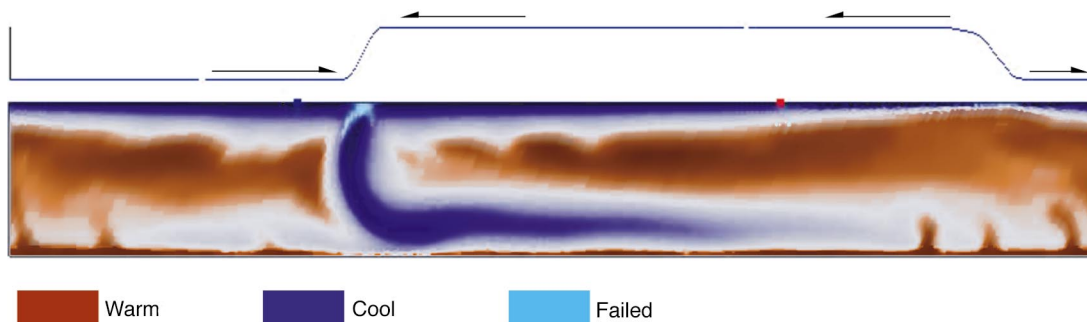


Figure 3. This shows convection in a system with Lagrangian particle tracking that allows the use of a strain-weakening yield stress. Dark blue areas are cool, red areas warm. Viscosity increases strongly as the material cools. The light-blue region indicates that yielding at a high strain rate takes place at the region where the cold material turns back down into the interior. The surface velocity plot shows that the motion is strongly partitioned into broad, nondeforming regions separated by narrow zones of deformation: Typical velocities are a few millimeters to a few centimeters per year. Note how the zone of yielding has changed from Figure 2c; it is now considerably more localized and focused into a single band that separates the two converging regions.

the point of failure does not accumulate—missing what we consider to be one of the most important features of Earth's lithosphere.

To directly incorporate such effects requires the advection of the strain history through the mesh. Even the simplest formulations therefore encounter the advection–interpolation difficulty, which rapidly degrades the stored information. One alternative is to parameterize the rheological weakening associated with the strain history while ignoring the advection of the history—producing an effective rheology with strain-rate weakening.⁵ However, this only partially solves the need for a damage model. The simplest complete solution would be to introduce a Lagrangian framework for the strain variables.

One approach we recently tried is using the tracer-particle scheme to track the strain history. The tracer particles can carry any amount of information, including tensor quantities. Their values are updated from the mesh-based variables by standard finite-element interpolation from the shape functions. The Lagrangian reference frame ties back into the mesh by using the particle locations in place of the usual quadrature points for computing element integrals. The method comes from recent work to extend the particle-in-cell schemes used in the 1950s and 1960s.⁶ Some subtleties are required for modeling convecting fluids associated with the enormous strains, but the method is essentially as simple as outlined earlier.

Figure 3 shows the result of such a calculation. The grid resolution is similar to that in Figure 2, but now there are also approximately 8 to 12

particles in every element that store the material's deformation and yielding history.

The rheological law is, once again, about as simple as can be. The yield stress in the material decreases as a function of the amount of strain the material accumulates while yielding takes place (to a total of about half the initial strength). This produces a feedback effect in which strain and strain rate localize very strongly and dominate the system's subsequent evolution. The localized zone close to the downwelling is light blue; it forms a linear feature that becomes approximately locked in place. Despite the relatively high resolution and the restriction to 2D, the localization is still far too indistinct to possess the properties of a genuine fault.

The time evolution of the simulations shown in Figures 2 and 3 shows very clearly how subtle changes in the material properties of the lithosphere can produce dramatically different behavior.

The spherical Earth

The spherical shell geometry for Earth's mantle has an important influence on the convection pattern in map view and must be treated realistically for global problems. We also must abandon flat-Earth models when modeling anything other than regional-scale observable data, including plate motions, the gravitational field, Earth's rotation and polar wander, and mantle seismic structure—all global features.

Several recent studies have focused on going beyond the flat-Earth approximation.⁷ Model-

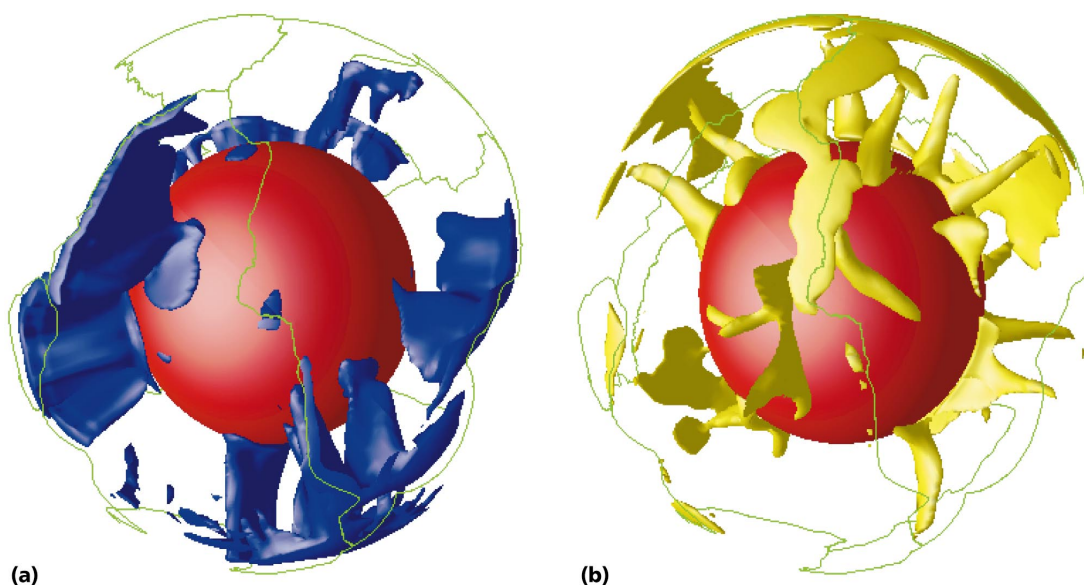


Figure 4. This full-spherical shell model of mantle convection shows (a) low-temperature isosurfaces in blue, (b) high-temperature isosurfaces in yellow, and the core-mantle boundary in red. The viscosity depends on both temperature (the cold areas being strong) and position (plate margins being weak). The green lines represent present-day plate margins. The cold downwellings have a distinct sheet morphology reminiscent of the structures imaged by seismic tomography, whereas the hot upwellings have a more complex, plume-like morphology at mid-depth in the simulation.

ing mantle dynamics in a spherical shell geometry ultimately has two important goals. The first goal is to investigate the planform of mantle convection with realistic rheology and geometry and to understand the nature of seismically observed mantle structure. The second is to explain surface observables such as plate motion.

Of necessity, whole Earth simulations are enormously demanding computationally and require the most efficient computational grids, numerical algorithms, and parallel computing. The simulations we present here are the product of a computational tool that Shijie Zhong and his colleagues developed.⁸ One benefit of choosing an FEM is that the bones of the spherical code are essentially identical to those of the flat-Earth code, and the same efficient solution algorithms (full multigrid, Figure 1c3) can be applied. The mesh comprises 12 logically rectangular sub-meshes that are used to tile the sphere uniformly. The rectangular meshes are fully subdividable to allow an efficient implementation of the multigrid algorithm.

Parallel computation is implemented using a message-passing algorithm on the structurally enforced communication (between the 12 subgrids) and, if enough processors are available, within each of the subgrids. The full multigrid algorithm resembles the one described earlier but with special features that result from the re-

quirement of the grid and parallel computing.

As an example, we present a time-dependent thermal convection calculation with tectonic plates, temperature-dependent rheology and layered viscosity, and predominantly internal heating.⁸ Combining temperature-dependent and reduced viscosity at plate margins results in mobile surface plates. The reduced viscosity at plate margins simulates the effects of enhanced deformation associated with interplate seismicity.

The model starts with a small amount of negatively buoyant (cold) material in subduction zones. The model is then integrated for about two turnover times so that the heat transfer and flow field are statistically time-invariant.

Figure 4 shows a snapshot of the characteristic thermal structure from this calculation. Figure 4a shows how the cold downwellings originate at plate margins and sink into the mantle as sheet structures. The downwellings impinge the core-mantle boundary and spread over the boundary. The sheet structure resembles the morphology of Benioff zones of deep seismicity as well as tomographic images of seismic velocity below the termination of seismicity at the 660-km depth.⁹

As shown in Figure 4b, throughout most of the mantle, the upwellings have plume-like morphologies. Near the surface, upwellings are long and linear and are closely associated with diverging plate boundaries. The hot plumes com-

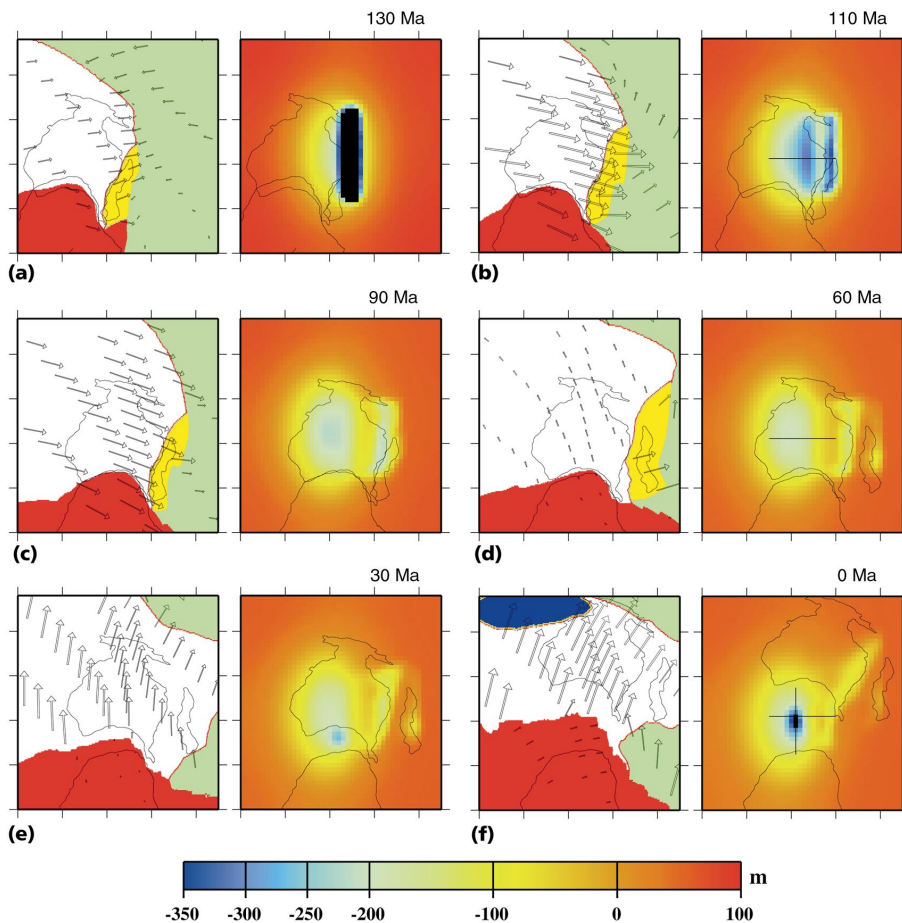


Figure 5. Figures a through f show the results of our 3D time-dependent model of thermal convection in the Australian region from 130 million years ago to the present. The left box for each time frame shows the tectonic plates and their velocities, which are imposed as kinematic boundary conditions. Inside the domain (not shown), the full dynamic coupling of Equations 1 to 3 are solved. The right box for each time frame shows the resulting dynamic topography resulting from the fluid flow. It matches the observed changes in the shape of Earth, inferred from sea-level change, extremely well. (Figure courtesy of the American Association for the Advancement of Science.)

mence at the core-mantle boundary and are preferentially located below either spreading centers or plate interiors, including those in the Pacific. The hot plumes, when impinging the surface, might produce extensive volcanic activity and might be responsible for volcanic chains such as Hawaii and Iceland.

By controlling the locations of downwelling structures, surface plates determine the scale of thermal structure within the mantle. The downwelling sheet and upwelling plume structures and their spatial scales are consistent with the observations of seismic structures from seismic tomography. Global models such as this can also relate interior dynamics to surface plate motion and the geoid, and therefore impose important constraints on how mantle rheology depends on temperature.¹⁰

Australian case study

A fundamental but poorly constrained aspect of mantle convection is its time dependence. This is especially unfortunate, because we know that plate kinematics change substantially

through time on a variety of space and time scales. Moreover, nearly all models of mantle convection, including those shown in Figures 2 and 4, are time dependent, at least for part of their history.

Because convection is a buoyancy-driven phenomenon, it would be beneficial to find other observations that are also sensitive to mantle flow. Topography and gravity, which are sensitive to internal mass anomalies, provide a powerful constraint on mantle dynamics, as do tomographic images of the mantle constructed from seismic waves. Unfortunately, such observations only provide us with present-day, instantaneous snapshots of convection. However, buoyancy-driven flows cause deflections of density interfaces, a contribution to surface topography called *dynamic topography*. Such deflections of Earth's surface cause relative sea-level fluctuations that leave a decipherable record in the geological history of the continents.

To understand the vertical motion of mass within the mantle as a function of time, we integrated a variety of geological observations in the Australian region with the aid of time-dependent 3D mantle convection models.¹¹ We have focused

on the Australian region because large-scale features exist there that have never been adequately explained by the application of the kinematic rules of plate tectonics but that might be associated with the radial motion of buoyancy within the mantle.

We focused on the subsidence and then emergence of the Australian continent out of the sea during the Cretaceous Period, which was a period of extensive marine flooding of the other continents worldwide by a sea-level rise. We also focused on the existence of a present-day cold spot, often interpreted as a convective downwelling, in the mantle below the spreading center between Australia and Antarctica (a region known as the Australian–Antarctic discordance, or AAD). We have shown that the two features are related to the overriding of a downwelling (a long-lived subduction zone) by the Australian plate.

We built a 3D model of the mantle around Australia using the numerical techniques that we have just described. Figure 5 shows the results. One of our most important geological constraints is the observed kinematic history of the plates (referred to as paleogeography) over the last 100 million years or so. In principle, we could develop a dynamic model with the aim of reproducing observed plate motions. However, this problem is highly nonlinear; our understanding of the feedback between plate margin rheology and plate motions is still in its infancy, and the computational cost is too formidable for us to realistically pursue this approach at present. Instead, we imposed observed plate motions with velocity boundary conditions onto a 3D domain while making the flat-Earth assumption described earlier.

Starting with Figure 5a, the models show a slab—a convective downwelling—about 1,400 km from the restored eastern margin. The models suggest that above cold fluid, the viscous stresses pulled the surface downward so that a deep trough developed off the coast of Australia. Figures 5a through 5c show that, as Australia moved east in a fixed reference frame and over the cold fluid from 130 to 190 million years ago, a broad dynamic topography depression of decreasing amplitude migrated west across the continent. This caused the continent to subside and then uplift, as shown in Figures 5c through 5e. During this period, most of the slab or cold fluid descended into the deeper mantle, but the models show that part of the cooler mantle became trapped within the so-called transition zone (a region between 410 and 660 km deep, bounded by two phase transitions).

From 40 million years ago to the present, wisps

of this cool mantle were drawn up by the northwardly migrating diverging plate boundary between Australia and Antarctica (as shown in Figures 5e through 5f). This caused a circular dynamic topography depression to develop at the present position of the AAD. Using the observations to constrain how topography changes as a function of time, we can follow the fall and rise of mass within the mantle over a period of about 100 million years.

The implication of this regionally focused study is that disparate data types can be understood within the context of time-dependent mantle convection when convection models are explicitly tailored to observed paleogeography. As convection models begin to incorporate multiphysics, there seems to be no limit to the manner in which geological observations, which constrain the time domain, can be integrated with present-day geophysical observations.

Our goal has been to combine the Stokes and energy equations with a realistic rheology, thereby letting us understand the complex dynamic coupling that occurs in the mantle and that gives rise to plate tectonics and other surface features. This approach holds great promise because it makes a tremendous amount of data relevant to understanding Earth's dynamics. The challenge is that the computational models must be inherently realistic, particularly when predicting observed geography or plate history, so that the models can be connected with observations. We view this as one of the most exciting future directions of computational geodynamics.

This raises an important philosophical issue that many areas of computational science face: Should a model attempt an explicit match with observations, or is a match in a statistical sense sufficient? In the case of geophysics, we must ask whether we wish to produce a model of Earth as it really is today or of an Earth that might have been if the chaotic processes of formation and evolution happened to work out differently.

Clearly, one important goal is to determine whether a fully dynamic model can predict the kinematic rules of plate tectonics. That is, can we create a generic model of Earth that looks correct—a model in which plates have about the correct size distribution, have the same degree of sharpness around their edges, and move at about the same velocity as observed? Obviously, the acid test regarding whether such generic models really do look correct is whether they can predict the spe-

cific information that we do have about the evolution of our particular “possible Earth.” Hence, our conviction that integrated observation and modeling-driven research is of fundamental importance.

Important developments must occur if we are to achieve still more realistic matches with observations: First, in the realm of plate-margin and shear-zone physics, for example, we must incorporate explicit, history-dependent rheologies driven by changes in the grain size or incorporation of volatiles (if indeed this is the relevant controlling physics) to better simulate plate tectonics. Initially this can be achieved through a flat-Earth approximation in two dimensions. Eventually this will have to be extended into a full 3D spherical geometry, although some of the fine-scale physics will inevitably need to be parameterized, just as it is in general circulation models of Earth’s atmosphere and ocean system.

Second, because it is not yet clear which observations best constrain the dynamics of plate boundaries, predictions from geodynamic models should guide the acquisition of the most relevant observations on plate boundaries.

Third, in a problem that areas of computational science face as more complex physics is coupled into codes, the software will have to undergo greater levels of benchmarking to prove its veracity. Indeed, computational geodynamics has been placed on a much firmer footing since the first benchmark comparisons of codes a decade ago.

Finally, it is not yet possible to fully compute mantle convection in a sphere with the coupled physics that gives rise to plate tectonics. However, given the newly developed finite-element codes, such models will achieve sufficient resolution with an order-of-magnitude increase in CPU speeds and memory over the largest massively parallel supercomputers at national and international centers. ■

Acknowledgments

The Australian Geodynamics Cooperative Research Centre (AGCRC) partially funded Louis Moresi’s work, and the AGCRC’s director has given permission for this article’s publication. This article is contribution number 8676 of the Division of Geological and Planetary Sciences, Caltech. This work is funded in part by the National Science Foundation and the National Aeronautics and Space Administration.

References

1. P. Kearey and F.J. Vine, *Global Tectonics*, 2nd ed., Blackwell Science Ltd., London, 1996.
2. W.L. Briggs, *A Multigrid Tutorial*, Society for Industrial and Applied Mathematics, Philadelphia, 1987.
3. U. Christensen, “Convection with Pressure and Temperature Dependent on Newtonian Rheology,” *Geophysical J.*, Vol. 77, 1984, pp. 343–384.
4. L. Moresi and V. Solomatov, “Mantle Convection with a Brittle Lithosphere: Thoughts on the Global Tectonics Styles of the Earth and Venus,” *Geophysical J. Int’l*, Vol. 133, No. 3, June 1998, pp. 669–682.
5. P. Tackley, “Self-Consistent Generation of Tectonic Plates in Three-Dimensional Mantle Convection,” *Earth & Planetary Science Letters*, Vol. 157, Nos. 1–2, Apr. 1998, pp. 9–22.
6. D. Sulsky, Z. Chen, and H.L. Schreyer, “A Particle Method for History Dependent Materials,” *Computer Methods in Applied Mechanics and Eng.*, Vol. 118, Nos. 1–2, Sept. 1994, pp. 179–196.
7. P.J. Tackley et al., “Effects of an Endothermic Phase-Transition at 670-km Depth in a Spherical Model of Convection in the Earth’s Mantle,” *Nature*, Vol. 361, No. 6414, Feb. 1993, pp. 699–704.
8. S. Zhong et al., “The Role of Temperature-Dependent Viscosity and Surface Plates in Spherical Shell Models of Mantle Convection,” to be published in *J. Geophysical Research*.
9. S.P. Grand, R.D. van der Hilst, and S. Widiyantoro, “Global Seismic Tomography: A Snapshot of Convection in the Earth,” *GSA Today*, Vol. 7, No. 4, Apr. 1997, pp. 1–7.
10. S. Zhong and G.F. Davies, “Effects of Plate and Slab Viscosities on the Geoid,” *Earth & Planetary Science Letters*, Vol. 170, No. 4, July 1999, pp. 487–496.
11. M. Gurnis, R.D. Müller, and L. Moresi, “Dynamics of Cretaceous Vertical Motion of Australia and the Australian-Antarctic Discordance,” *Science*, Vol. 279, No. 5356, Mar. 1998, pp. 1499–1504.

Louis Moresi is a research scientist with the CSIRO Division of Exploration and Mining in Perth, Australia. He develops finite-element methods for the study of highly deforming Earth materials and applies the codes to the study of continental deformation, plate tectonics, convection, and heat flow. He has a DPhil from Oxford University in geophysics and a BA from Cambridge University in physics. Contact him at CSIRO Exploration & Mining, PO Box 437, Nedlands, 6009, Western Australia; l.moresi@ned.dem.csiro.au.

Michael Gurnis is a professor of geophysics in the Seismological Laboratory at Caltech. His research interests are in the application of computational mechanics to problems related to plate tectonics and the Earth’s deep interior. He is a fellow of the American Geophysical Union and a senior fellow of the Geological Society of America. He has a PhD in geophysics from the Australian National University. Contact him at MS 252-21, Caltech, Pasadena, CA 91125; gurnis@caltech.edu.

Shijie Zhong is a research scientist at MIT. His research involves the use of large-scale computer models to study the deep interiors of the Earth and other terrestrial planets. He earned a BS in geophysics from the University of Science and Technology of China and a PhD in geophysics and scientific computing from the University of Michigan. He is a member of the American Geophysical Union. Contact him at the Dept. of Earth, Atmospheric, and Planetary Sciences, MIT, Bldg. E54-511, Cambridge, MA 02139; szhong@rayleigh.mit.edu.